9/623068 525 Rec'd PCT/PTO 28 AUG 2000

Cover Letter and Mailing Certificate for submission of additional Oath for PCT/US99/04376 upon 30 month entry into th US National Stage

5028

This additional original oath is submitted for inclusion with PCT/US99/04376 upon 30 month entry into the U.S. National Stage. This additional original oath refers to amendments requested to be made to the application upon entry into the National Stage. A letter requesting amendments and replacement pages to effect these amendments are enclosed herewith. The amendments are considered by the inventors to add **no new matter to the specification.**

The 30 month deadline for 30 month entry into the U.S. National Stage is August 26, 2000, a Saturday. These papers are being submitted on August 28, 2000, the first business day following August 26, 2000. Thus these papers are considered by the inventors to be submitted at entry into the U.S. National Stage.

A transmittal letter for entry into the National Stage with appropriate fees, oath, small entity statements, replacement pages, copy of Statement under Article 19 (1) and other documents were previously filed for PCT/US99/04376 by Express Mail under Express Mail Label EK922461352US on August 26, 2000.

The following items have been deposited with the U.S. Postal Service today, August 28, 2000 and sent by Express Mail to Assistant Commissioner for Patents, Box PCT, Washington, D.C. 20231. Express Mail Label:

Enclosed are: EL049304271US

1) Signed and filled out declaration PTO/SB/01, 2 pages. (original)
2) Signed and filled out declaration additional inventors PTO/SB/02A, 1 page. (original)
This declaration or declarations (or oaths) refer specifically to amendments, including requested amendments.
3) Signed 2 page letter requesting amendments that do not add new matter to the specification.
4) Replacement pages 5, 6, 7, 8, 38, 43, 46; a total of 7 replacement pages with header "PCT/US99/04376(Entry U.S. National Stage Aug. 2000)"all on size A4 paper.
5)Copy of Statement under Article 19(1) on size A4 paper.
6)Return receipt postcard.

Deposited by Robert McGinnis with U.S. mail today, August 26, 2000.
Respectfully submitted,

Robert McGinnis

Robert McGinnis, Patent Agent 44, 232

1/2 **<u>Letter requesting Amendments t PCT Applicati n PCT/US99/04376 up n 30 month entry into th U.S. Nati nal Stage (August 26, 2000)"</u>**

(The amendments do not add new matter to the specification. Claims are being cancelled.)

The applicants hereby request the following amendments to the application upon entry into the U.S. National Stage:

<u>Amendments to the background of the application:</u>

1) page 5 line 22 after "territory" insert --and that it was difficult to predict the power of using a less dense map at that time--

2) page 5 line 22 after "[10]" insert -- **The inventor's work, however, is a predictor of the power and success of a less dense map.**--

A replacement page 5 with header "PCT/US99/04376(U.S. National Stage Entry Aug. 2000)" is enclosed to effect these amendments 1) and 2) to the background.

3) page 6 line 22 after "TDT," change the text "***to increase the likelihood of conditions occurring that increase the power of the TDT in the linkage study, the bi-allelic markers used in the study are chosen so that the least common allele frequencies of the markers vary systematically over a range or subrange of least common allele frequency.***" from bold face italics to regular italics with underlining. A replacement page 6 with header "PCT/US99/04376(U.S. National Stage Entry Aug. 2000)" is enclosed to effect amendment 3) to the background.

4) page 7 line 6 after "TDT," change the text "***to increase the likelihood of both criteria (1) and (2) occurring for one or more markers, so as to increase the power of the TDT in the linkage study, the bi-allelic markers used in the study are chosen so that the least common allele frequencies of the markers vary systematically over a range or subrange of least common allele frequency AND the chromosomal location of the markers vary systematically over one or more chromosomes or chromosomal regions. And the bi-allelic markers are chosen so that the markers' chromosomal locations and least common allele frequencies vary systematically in an essentially independent manner.***" from bold face italics to regular italics with underlining.

5) page 7 line 32 delete the text in brackets [***In addition, the two-dimensional linkage study techniques do not necessarily favor using markers in a scan that are about evenly spaced along a chromosome as in the conventional techniques.*** *This is because*]. On page 7 line 31 after the text "unfavorably" insert --**Conventional techniques use a one-dimensional concept of "closeness". *These techniques space markers about evenly along a chromosome in the hope that some markers will be "close" (on the chromosome) to the sought gene. (They also favor bi-allelic markers with least common allele frequencies near 0.5.) These*--**.

A replacement page 7 with header "PCT/US99/04376(U.S. National Stage Entry Aug. 2000)" is enclosed to effect amendments 4) and 5) to the background.

6) page 8 line 18 insert on the next line after "background" -- **<u>Summary</u>**

Versions of the invention use a new, two-dimensional concept of "closeness" for association-based linkage studies. Versions of the invention use bi-allelic markers that "cover" or are distributed approximately evenly (or systematically) over two-dimensional regions. These regions have the two dimensions of chromosomal location and least common allele frequency. **Conventional techniques suffer from a kind of one-dimensional lack of depth perception.** (They also favor bi-allelic markers with least common allele frequencies near 0.5.) Two-dimensional linkage study techniques overcome this lack of depth perception. These two-dimensional techniques greatly increase the chance that one or more markers used in a study will be close to the sought gene in two-dimensions. This results in more powerful, systematic and efficient methods (including computer programs) and machines for finding genes, such as harmful

genes and genes of only modest effect. These techniques also use less dense (more efficient) marker maps (or marker "coverings").

The basic principles behind the two-dimensional approach spawn numerous other inventions. These include methods, machines and compositions of matter (groups of molecules) used for gathering the data (i.e. genotype/sample allele frequency data) used in the new two-dimensional studies, and computer techniques for using and handling such data. These techniques work for creatures other than human beings. And they work for markers and genes that are not bi-allelic (any marker or gene can be mathematically transformed to behave like it is bi-allelic). This summary is not exhaustive or limiting, there are other inventions not listed or specifically described here.--

A replacement page 8 with header "PCT/US99/04376(U.S. National Stage Entry Aug. 2000)" is enclosed to effect amendment 6) to the background.

Amendments to the Description

7) page 38 line 2, page 38 line 17 and page 38 line line 20 delete the text "Best Mode" and replace the text "Best Mode" with the text "Set/Subset Example". A replacement page 38 with header "PCT/US99/04376(U.S. National Stage Entry Aug. 2000)" is enclosed to effect the amendments to the description under item 7).

8) page 43 line 4 delete the text "Best Mode" and replace the text "Best Mode" with the text "Set/Subset Example". A replacement page 43 with header "PCT/US99/04376(U.S. National Stage Entry Aug. 2000)" is enclosed to effect the amendment to the description under item 8).

9) page 46 line 24 and line 28 delete the text "Best Mode" and replace the text "Best Mode" with the text "Set/Subset Example". One page 46 lines 26 and 27 delete the text "Best Mode" and replace the text "Best Mode" with the text "Set/Subset". A replacement page 46 with header "PCT/US99/04376(U.S. National Stage Entry Aug. 2000)" is enclosed to effect the amendments to the description under item 9).

Canceling of Claims and presentation of uncancelled claims for examination

The applicants hereby request that all claims in the application be cancelled except for the following claims that were filed April 17, 2000: Claims 3, 4, 5, 7, 8, 20, 21, 22, 23, 33, 34, 35, 37, 38, 50, 51, 52, 53, 54, 57. Thus the applicants request that only claims 3, 4, 5, 7, 8, 20, 21, 22, 23, 33, 34, 35, 37, 38, 50, 51, 52, 53, 54, 57 filed April 17 2000 be examined.

I hereby attest that no new matter is added to the specification of the application by the amendments requested in the two pages of this letter.

Respectfully submitted,

Robert McGinnis

Robert McGinnis
U.S. Patent Agent 44, 232

0.5/0.5. Secondly, bi-allelic markers with lower least common allele frequencies, less than 0.3(0.7/0.3) or 0.2(0.8/0.2), are viewed unfavorably for linkage studies in this reference. Thirdly, the early version of the criterion of "information content" of markers used in this reference was based on sib pair analysis and the later, current version of the criterion, does not depend on any particular test for linkage.[5, 6] **Thus, the criterion of information content in this reference, has never specifically employed the TDT (transmission disequilibrium test) or any association based test, whereas the two-dimensional linkage study techniques of this application are based on a completely different perspective of using association based tests.** (This reference[4] is not admitted to be prior art with respect to the present invention by it's mention in this background.)

<u>Increased Power of the TDT (transmission disequilibrium test)</u>

Characteristics of a new type of linkage test, the TDT (transmission disequilibrium test), were described in 1993. The inventor, R.E.McGinnis, was one of the authors of this reference.[7] In 1996, Risch and Merikangas argued that conventional linkage analysis has limited power to detect genes of modest effect. And Risch and Merikangas attempted to illustrate the increased power of association based linkage tests such as the TDT over other types of conventional linkage tests.[8] However, Risch and Merikangas' analysis was criticized by Muller-Myhsok and Abel as being based on the optimal assumption that the analyzed allele was the disease allele itself. Muller-Myhsok and Abel concluded that researchers should be aware that the power of association studies such as the TDT can be greatly diminished in more common, less optimal situations.[9] In their response to Muller-Myshok and Abels' letter, Risch and Merikangas essentially agreed with the logic of Muller-Myshok and Abels' criticism. Risch and Merikangas stated that to a large extent, the expectation with respect to linkage disequilibrium across the genome is uncharted territory and that it was difficult to predict the power of using a less dense map at that time.[10] **The inventor's work, however, is a predictor of the power and success of a less dense map.** (None of the references in this paragraph [7,8, 9,10] is admitted to being prior art with respect to the present invention by their mention in this background.)

<u>More Detailed Studies of the Power of the TDT</u>

The inventor, R.E.McGinnis, has done extensive investigations on the power of the TDT. His observations and calculations of the increased power of the TDT in many situations have been

---

[5] Kruglyak, et. al.: Complete Multipoint Sib-Pair Analysis of Qualitative and Quantitative Traits. Am J Hum Genet, 1995, vol. 57: pp. 439-454.

[6] Kruglyak, et. al.: Parametric and Nonparametric Linkage Analysis: A Unified Multipoint Approach. Am J Hum Genet , 1996, vol. 58, pp. 1347- 1363.

[7] Spielman, R.S., McGinnis, R.E., Ewens, W.J.: Transmission Test for Linkage Disequilibrium: The Insulin Gene Region and Insulin-dependent Diabetes Mellitus(IDDM). Am J Hum Genet, 1993, vol. 52, pp. 506-516.

[8] Risch, N. and Merikangas, K.: The Future of Genetic Studies of Complex Human Diseases. Science, 13 September 1996, vol. 273, pp. 1516-1517.

[9] Muller-Myshok, B. and Abel, L.: Technical Comments: The Future of Complex Diseases. Science, 28 February 1997, vol. 275, pp. 1328-1329.

[10] Risch, N. and Merikangas, K.: Technical Comments: The Future of Complex Diseases. Science, 28 February 1997, vol. 275, p. 1330.

published.[11] In this paper a general framework for determining the power of the TDT in many different situations is presented. The analysis of Risch and Merikangas [8] and others is shown by the inventor to be a special case of his general framework. His observations and calculations published in this paper have shown that the TDT has increased power in more common, less optimal situations as well as the less common, optimal situation cited by Muller-Myshok and Abel [9]. As opposed to the observation of Muller-Myhsok and Abel, the inventor's calculations indicate that association tests such as the TDT have increased power in typical situations even when the ratio m/p departs significantly from unity and, or the linkage disequilibrium between the analyzed (marker) allele and disease polymorphism is only half its maximum possible value. The inventor arrived at these conclusions independently and did not derive them from others.

A Major Conclusion Drawn by the Inventor about the TDT and Linkage Studies: Using Bi-allelic Markers of Systematically Varying Allele Frequencies Increases the Power of Linkage Studies Using the TDT

The inventor's calculations and observations about the increased power of the TDT in more common, less optimal situations led him to the conclusion that the power of linkage studies using the TDT is greatly increased under some conditions. Under some conditions, the power of the TDT in a linkage study using bi-allelic markers is greatly increased when each of one or more of the bi-allelic markers used in the study fulfill two criteria: (1) the allele frequencies of each of the one or more of the bi-allelic markers are similar (but not necessarily the same, or even approximately the same) as the allele frequencies of an unknown bi-allelic gene causing a disease under study; and (2) each of the one or more bi-allelic markers is in some degree of linkage disequilibrium with the gene. Thus for a typical linkage study using bi-allelic markers and the TDT, *to increase the likelihood of conditions occurring that increase the power of the TDT in the linkage study, the bi-allelic markers used in the study are chosen so that the least common allele frequencies of the markers vary systematically over a range or subrange of least common allele frequency.* This major conclusion of the inventor's research is quoted directly from his unpublished manuscript that was included with previously filed U.S. Provisional Patent Applications:"This example is typical and highlights perhaps the most important finding of this paper; namely the importance of using bi-allelic markers with heterozygosity similar to that of a bi-allelic disease gene. Indeed, since a majority of susceptibility loci may be bi-allelic, the judicious use of bi-allelic markers of both high, medium and low heterozygosity may be crucial in order to detect and replicate linkages to loci conferring modest disease risk." (page 25) (In this context the phrase "bi-allelic markers with heterozygosity similar to that of a bi-allelic disease gene" is essentially equivalent to "bi-allelic markers with individual allele frequencies similar to those of a bi-allelic disease gene" and "bi-allelic markers of both high, medium and low heterozygosity " is essentially equivalent to the phrase "bi-allelic markers whose least common individual allele frequencies are high, medium and low".)

Systematically Varying Both Marker Chromosomal Location and Marker Allele Frequency of Markers in Linkage Studies

---

[11] McGinnis, R.E.: Hidden Linkage: Comparison of the affected sib pair (ASP) test and transmission disequilibrium test (TDT). Annals of Human Genetics, 1998, vol. 62, pp. 159-179.

The inventor's calculations and observations have demonstrated the increased power of the TDT in more common, less optimal situations when a bi-allelic marker and bi-allelic gene have (1) similar but not identical allele frequencies and (2) the marker and gene are in some degree of linkage disequilibrium. Thus, for a typical linkage study using bi- allelic markers and the TDT, *to increase the likelihood of both criteria (1) and (2) occurring for one or more markers, so as to increase the power of the TDT in the linkage study, the bi-allelic markers used in the study are chosen so that the least common allele frequencies of the markers vary systematically over a range or subrange of least common allele frequency AND the chromosomal location of the markers vary systematically over one or more chromosomes or chromosomal regions. And the bi-allelic markers are chosen so that the markers' chromosomal locations and least common allele frequencies vary systematically in an essentially independent manner.*

Two-dimensional Linkage Study Techniques

As has been stated, conventional linkage study scanning techniques use markers that are distributed approximately evenly in the dimension of chromosomal location. These conventional, one dimensional, scanning techniques focus primarily on the chromosomal location of markers used in a scan and give little attention to the dimension of allele frequency.[1, 2, 3]

One of the main implications of the inventor's work is to use a set of bi-allelic markers for a typical linkage study using the TDT (or other association-based linkage test) wherein the chromosomal locations and least common allele frequencies of the markers in the set systematically vary in an essentially independent manner over the dimensions of chromosomal location and least common allele frequency respectively. This is equivalent to using a set of bi-allelic markers for a linkage study scan wherein the set of markers systematically scan or "cover" a two-dimensional region having dimensions of chromosomal location and least common allele frequency. (Such a two-dimensional region can be thought of as an area in an x-y plot or a group of squares on a chessboard.)

In addition, the inventor's calculations and observations indicate that bi-allelic markers having least common allele frequencies less than 0.3, 0.2 or even less than 0.1 have an important place in linkage studies using association based linkage tests. This is markedly different than Kruglyak's information content evaluation of bi-allelic markers for use in linkage studies, in which bi-allelic markers with least common allele frequencies less than 0.3 or 0.2 are viewed unfavorably.[4]

Conventional techniques use a one-dimensional concept of "closeness". ***These techniques space markers about evenly along a chromosome in the hope that some markers will be "close" (on the chromosome) to the sought gene. (They also favor bi-allelic markers with least common allele frequencies near 0.5.) These conventional techniques suffer from a kind of one dimensional view or lack of depth perception.*** **In the conventional techniques, a marker can look very close to a gene's location in terms of chromosomal location, but the marker can be very far from the gene's location in the new two-dimensional view used by versions of the invention. It is as if the conventional 1D techniques look at a chessboard from on edge. Markers and a gene which are on different squares of the board, but in the same column of squares, look very close to each other when the board is looked at from on edge. But when the board is looked at**

from the top in 2D, two dimensions, markers which looked very close to each ther and the gene before (when looking from on edge) can be seen to be very far from the gene.

<u>Further Implications of the Two-dimensional Linkage Study Perspective</u>

These two-dimensional techniques work when multiple genes cause a genetic characteristic and are effective in searching for these genes. A two-dimensional bi-allelic marker "covering" or scanning approach also increases the power of linkage studies using other association based linkage tests such as the AFBACmethod, the haplotype relative risk (HRR) method [12], and comparison of marker allele frequencies in disease cases and unrelated controls[13]. These references[12, 13] are not admitted to being prior art with respect to the present invention by their mention in this background.)

<u>Patents That May Be Helpful In Starting A Search Of The Background</u>

Some patents that are in the same general areas as versions of the invention are cited here: US Patent Number 5,667,976 Solid supports for nucleic acid hybridization assays. Published International Application WO 98/20165 Biallelic Markers. Published International Application WO 98/07887 Methods for treating bipolar mood disorder associated with markers on chromosome 18 p. US Patent Number 5,552,270 Methods of DNA sequencing by hybridization based on optimizing concentration of matrix-bound oligonucleotide and device for carrying out same. No patent in this paragraph is admitted to being prior art with respect to the present invention by it's mention in this background.

<u>Summary</u>

Versions of the invention use a new, two-dimensional concept of "closeness" for association-based linkage studies. Versions of the invention use bi-allelic markers that "cover" or are distributed approximately evenly (or systematically) over two-dimensional regions. These regions have the two dimensions of chromosomal location and least common allele frequency.

Conventional techniques suffer from a kind of one-dimensional lack of depth perception. (They also favor bi-allelic markers with least common allele frequencies near 0.5.) Two-dimensional linkage study techniques overcome this lack of depth perception. These two-dimensional techniques greatly increase the chance that one or more markers used in a study will be close to the sought gene in two-dimensions. This results in more powerful, systematic and efficient methods (including computer programs) and machines for finding genes, such as harmful genes and genes of only modest effect. These techniques also use less dense (more efficient) marker maps (or marker "coverings").

The basic principles behind the two-dimensional approach spawn numerous other inventions. These include methods, machines and compositions of matter (groups of molecules) used for gathering the data (i.e. genotype/sample allele frequency data) used in the new two-dimensional studies, and computer techniques for using and handling such data. These techniques work for creatures other than human beings. And they work for markers and genes that are not bi-allelic (any marker or gene can be mathematically transformed to behave like it is bi-allelic). This summary is not exhaustive or limiting, there are other inventions not listed or specifically described here.

---

[12] Falk CT and Rubenstein P: Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. Annals of Human Genetics, 1987, vol. 51, pp. 227-233.

[13] Bell GI, Horita S and Karam JH: A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. Diabetes, 1984, vol 33, pp. 176-183.

Versions of the apparatus comprise means for printing each of the one or more graphs.

Theory of Operation / Set/Subset Example

Systematically Varying Both Marker Chromosomal Location and Marker Allele Frequency of Markers in Linkage Studies

The inventor's calculations and observations have demonstrated the increased power of the TDT in more common, less optimal situations when a bi-allelic marker and bi-allelic gene have (1) similar but not identical allele frequencies and (2) the marker and gene are in some degree of linkage disequilibrium. Thus, for a typical linkage study using bi- allelic markers and an association based linkage test, ***to increase the likelihood of both criteria (1) and (2) occurring for one or more markers, so as to increase the power of an association based linkage test in a linkage study, the bi-allelic markers used in the study are chosen so that the least common allele frequencies of the markers vary systematically over a range or subrange of least common allele frequency AND the chromosomal location of the markers vary systematically over one or more chromosomes or chromosomal regions. And the bi-allelic markers are chosen so that the markers' chromosomal locations and least common allele frequencies vary systematically in an essentially independent manner.***

(In theTheory of Operation/ Set/Subset Example Section the traditional symbol used in scientific papers for the disequilibrium coefficient, δ, is used. This should not be confused with the symbol δ used for the covering distance in the remainder of the application. The symbol d is used for the disequilibrium coefficient in the sections of the application other than the Theory of Operation/Set/Subset Example Section.) The theory of operation is based on the mathematical observation that the TDT and other association-based tests for linkage are increased in power as the frequencies of the disease-causing allele of a bi-allelic gene and the positively associated allele of a linked bi-allelic marker become similar in magnitude. The inventor made this observation as a result of deriving the equation shown below for P_t (this is Equation 2 in the unpublished manuscript submitted for publication in December 1996 and in published paper by RE McGinnis in the Annals of Human Genetics vol 62, pp. 159-179, 1998).

$$P_t = .5 + (1 - 2\theta)\left[\frac{c_1c_4 - c_2c_3}{H}\right]\left\{p^2\left(\frac{\alpha^2 - \beta^2}{4}\right) + 2p(1-p)\left(\frac{(\alpha+\beta)^2 - (\beta+\gamma)^2}{16}\right) + (1-p)^2\left(\frac{\beta^2 - \gamma^2}{4}\right)\right\}$$

Equation 2

P_t may be regarded as the size of the "signal" which is given by the TDT to indicate that a tested marker is linked to a disease-causing gene. The more P_t is elevated above 0.5 (baseline), the greater is the evidence for linkage or "power" provided by the association-based linkage test known as the TDT.

Table 2 in the unpublished manuscript filed with previous US Provisional Patent Applications(see below) illustrates how signal strength increases substantially as the frequencies of disease-causing allele and positively associated marker allele become similar in magnitude. As noted on pages 24 and 25 of the unpublished manuscript(see below), Table 2 assumes that the frequency (p)

judicious use of bi-allelic markers of both high, medium, and low heterozygosity may be crucial in order to initially detect and replicate linkages to loci conferring modest disease risk.

Set/Subset Example:

Method for locating disease causing polymorphism using biallelic linkage analysis

Objective : To test, by association-based linkage analysis (e.g., by TDT), whether a disease-causing polymorphism is located on a particular chromosome (e.g., human chromosome 4) or within a particular subregion of that chromosome.

PART 1 - Steps in conducting the association-based linkage test

Step 1

To conduct the test, first divide the chromosome or subregion of interest into segments that are short enough that polymorphisms within each segment are likely to be in linkage disequilibrium with each other. The division of a chromosome or subregion of interest into "segments" is conceptual (*not* physical) and is based on chromosomal maps such as those provided by the Whitehead Institute or Marshfield Foundation for Biomedical Research. Although disequilibrium has been observed in Finnish populations between polymorphisms that are 7 to 10 centimorgans (cM) apart, the chromosomal segments for searching for disease-causing polymorphisms in more genetically heterogeneous populations should be less than 1 cM long (e.g., 250,000 base pairs long). These chromosomal segments might or might not overlap each other (i.e., share some of their length in common); but the set of chromosomal segments should completely cover the entire chromosome or entire subregion of interest, so that a disease-causing polymorphism located anywhere on the chromosome or anywhere in the subregion of interest will be detected by the test.

Step 2

It is well known that increased disequilibrium between a marker and linked disease locus increases evidence for linkage provided by association-based linkage tests such as the TDT. However, what has not been recognized is that the specific allele frequencies of the marker locus can also have an enormous impact on the strength of evidence for linkage. I

the nearly identical information with respect to their linkage and association with a third polymorphism such as a disease locus. Hence one of the two bi-allelic markers would provide no additional information and its inclusion in the subset would not increase the likelihood of detecting linkage and association to a nearby disease locus.

Therefore, bi-allelic markers belonging to the same chromosomal segment and subset should not only have similar allele frequencies, the δ value between *each pair* of bi-allelic markers in the same subset should be substantially less than $\delta_{max}= q-q^2$. This assures that every bi-allelic polymorphism belonging to the subset provides much new (i.e. non-redundant) information about linkage and association to any nearby bi-allelic disease locus; thus testing each bi-allelic marker in the subset would increase the likelihood of detecting linkage to a disease locus.

Step4: Test for linkage

To test for (association-based) linkage to a bi-allelic disease locus, each bi-allelic marker in each subset from each chromosomal segment is tested *individually* by using the TDT, AFBAC method or other family-based linkage test. To conduct these tests for a particular marker, members of nuclear families (most especially parents, and any children who manifest disease) are genotyped at the marker being tested and the genotypes are then evaluated according to the TDT, AFBAC method or other family-based linkage/association test (for description of TDT and AFBAC, see Spielman et al, Am J of Human Genetics 52:506-516 (1993) and Thomson, Am J Human Genetics 57:487-498 (1995)). Alternatively, linkage and association is tested for each marker in each subset from each segment by genotyping individuals with disease and related or unrelated normal controls at each marker to be tested. (End of set/subset example)

Further Information

(Step 3 is not essential for the operation or utility of this version of the invention. In this set/subset example, the least common allele frequency subrange 0.1 to 0.5 is used. In versions of the invention similar to the set/subset example, versions of the invention are operable and have utility for any subrange of the least common allele frequency range 0 to 0.5. In addition, rather than genotyping DNA from single individuals in step 4, in some versions of the invention each marker in each subset from each segment is tested for association with disease by evaluating DNA from pooled samples.)

Statement under Article 19(1)

PCT/US99/04376

Some of the amended claims make use of the phrase "conditional probability", such as claim 11. Some of the amended claims make use of the phrase "proportion of groups", such as claim 14. There are various techniques to calculate or estimate such a probability or such a proportion. These techniques include, but are not necessarily limited to, direct calculation, statistical estimates, and Monte Carlo estimation techniques. Powerful software is available for calculation and statistical estimation for data in matrix format or two-dimensional format. Some such software is available from Cytel Software Corporation, Cambridge, Massachusetts (example: Exact Logistic Regression: Theory and Examples, Mehta CR, Patel NR, Statistics in Medicine, vol 14, 2143-2160(1995). Another example is SAS (SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513, USA.; A handbook of statistical analyses using SAS by Brian S. Everitt and Geoff Der, Boca Raton, Fla. : Chapman & Hall/CRC, 1998.). A further example is MATLAB (The MathWorks, Inc. 3 Apple Hill Drive, Natick, Mass. U.S.A. 01760-2098; MATLAB primer by Kermit Sigmon, 4th ed. Boca Raton : CRC Press, c1994.) Statistical techniques include techniques for hypothesis testing, goodness-of-fit and others.

The degree of skill in the art in probability and statistics is great. Indeed the inventor's important equation (Equation 2, page 38) is an equation for P_t, wherein P_t is a binomial probability for parental allele 'transmission' which determines the magnitude of the TDT chi-square statistic. P_S (pages 40-42) is also a binomial probability that determines the magnitude of the ASP test statistic. (see Abstract and Paper: Annals of Human Genetics (1998), 62, 159-179. The abstract is available on the World Wide Web and Internet, including at the journal's website.) Skill in the use of computers in the art is also great (page 25).

Some claims, such as claims 11, 12, 13, 14 and others make use of the phrase "substantially the known set of bi-allelic markers". As pointed out in the description (page 25) information on bi-allelic markers can be gained from sources such as the Whitehead Institute or Marshfield Foundation for Biomedical Research. Similar sources of information on Single Nucleotide Polymorphisms can be obtained from sources given in SNP attack on complex traits, Nature Genetics, volume 20 no. 3, Nov 1998, pp. 217-218.

Some claims, such as claims 11, 12, 13, 14 and others make use of the term "marker type" or similar terminology. As stated in the description, a bi-allelic marker may be an SNP, a microsatellite marker, a bi-allelic marker equivalent formed from one or more true bi-alleic markers. "Marker type" means type of true bi-allelic marker as for example an SNP or a microsatellite; or "marker type" means a bi-allelic marker equivalent of a certain type, such as a bi-allelic marker equivalent formed only from one or more SNPs or a bi-allelic marker equivalent formed only from one or more microsatellites.)